GMBA 7098: Statistics and Data Analysis (Fall 2014)

Hypothesis testing (2)

Ling-Chieh Kung

Department of Information Management
National Taiwan University

November 24, 2014

Road map

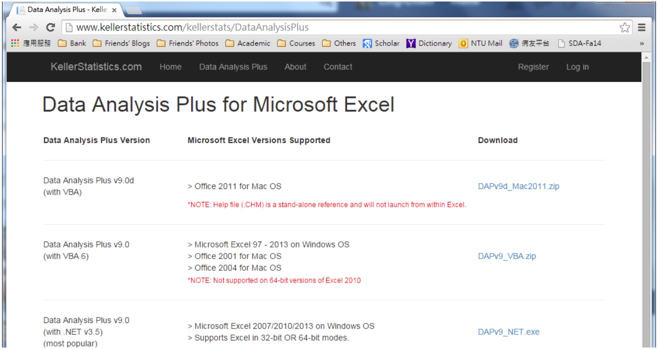
- **Preparations**.
- Testing population mean: variance known.
- Testing population mean: variance unknown.
- Testing population proportion.

Steps of hypothesis testing

- In conducting a test, write the following three parts:
  - **Hypothesis**: H_0 and H_a.
  - **Test**: The test to apply.
  - **Calculation**: Statistics, critical values, and/or p-values obtained by software.
  - **Decision and implication**: Reject or do not reject H_0? What does that mean?
- While the calculation part requires arithmetic or software, it is the "easiest" part.
  - Writing the correct hypothesis is the most important.
  - Writing a good concluding statement is also critical.

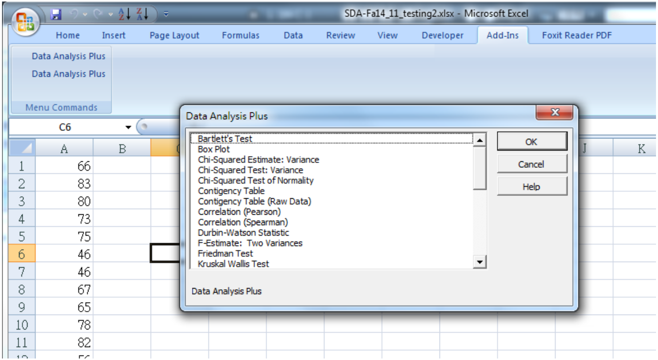
"Data Analysis Plus" (DAP)

- To do hypothesis testing by MS Excel, get "Data Analysis Plus" at `http://www.kellerstatistics.com/kellerstats/DataAnalysisPlus`.



"Data Analysis Plus" (DAP)

- Unzip it, double click the Excel file, and then open your own Excel files.
- Click "Add-Ins" and then "Data Analysis Plus:"



Road map

- Preparations.
- **Testing population mean: variance known**.
- Testing population mean: variance unknown.
- Testing population proportion.

Testing the population mean

- There are many situations to test the **population mean** μ.
  - Is the average monthly salary of fresh college graduates above \$22,000 (22K)?
  - Is the average thickness of a plastic bottle 2.4 mm?
  - Is the average age of consumers of a restaurant below 40?
  - Is the average amount of time spent on information system projects above six months?
- We will use hypothesis testing to test the population mean.
- Main factor:
  - Whether the **population variance** σ^2 is known.
  - Whether the population is normal.
  - Whether the sample size is large.

Testing the population mean

- When the population variance σ^2 is know:
  - If the population is normal or the sample size $n \geq 30$: z **test**.
  - In R: `z.test(x, alternative, mu, sigma.x, conf.level)`.[1]
  - In MS Excel: DAP $\rightarrow$ Z-Test: Mean.[2]
- When the population variance σ^2 is unknown:
  - If the population is normal or the sample size $n \geq 30$: t **test**.
  - In R: `t.test(x, alternative, mu, sigma.x, conf.level)`.
  - In MS Excel: DAP $\rightarrow$ T-Test: Mean.[3]
- Otherwise: Nonparametric methods (beyond the scope of this course).

[1] Execute first `install.packages("BSDA")` and then `library("BSDA")`.
[2] Or the built-in `ZTEST(array, x, sigma)`.
[3] There is no built-in method in MS Excel.

Example 1

- A retail chain has been operated for many years.
- The average amount of money spent by a consumer is \$60.
- A new marketing policy has been proposed: Once a consumer spends \$70, she/he can get one credit. With ten credits, she/he can get one toy for free.
- After the new policy has been adopted for several months, the manager asks: Has the average amount of money spent by a consumer increased? Let $\alpha = 0.01$.
  - Let μ be the average expenditure (in \$) per consumer after the policy is adopted. Is $\mu > 60$?
  - The population standard deviation is \$16.

Example 1: hypothesis and test

- The hypothesis is

$$H_0 : \mu = 60$$
$$H_a : \mu > 60.$$

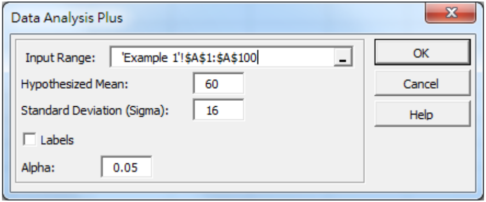
  - $\mu = 60$ is our **default position**.
  - We want to know whether the population mean **has increased**.
- Some researchers write

$$H_0 : \mu \leq 60$$
$$H_a : \mu > 60.$$

- Because the population variance is known and the sample size is large, we should use the z test.

Example 1: calculation

- The manager collects a sample with 100 purchasing records of consumers (in Sheet "Example 1" in "SDA-Fa14_11_testing2.xlsx.")
- In MS Excel: DAP → Z-Test: Mean. The **one-tailed** p-value is 0.0009.[4]



	A	B	C	D
1	**Z-Test: Mean**			
2				
3				*Column 1*
4	Mean			65
5	Standard Deviation			13.0601
6	Observations			100
7	Hypothesized Mean			60
8	SIGMA			16
9	z Stat			3.125
10	P(Z<=z) one-tail			0.0009
11	z Critical one-tail			1.6449
12	P(Z<=z) two-tail			0.0018
13	z Critical two-tail			1.96

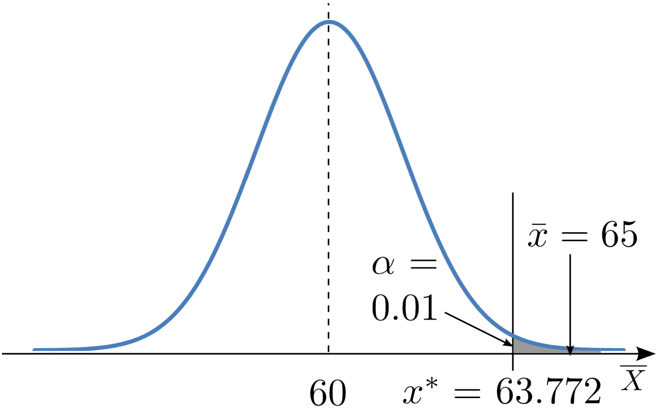
[4] In Excel, `ZTEST(A1:A100, 60, 16)` also gives 0.0009. In R, execute `z.test(x, alternative = "g", mu = 60, sigma.x = 16)`, where `x` is the vector containing the sample data.

Example 1: interpretation

- As p-value $= 0.000899 < 0.01 = \alpha$, we reject H_0.
- With a 99% confidence, the population mean is greater than 60.
- The new marketing policy ($70 for one credit and ten credits for one toy) is successful: Each consumer is willing to pay more (in expectation) under the new policy.

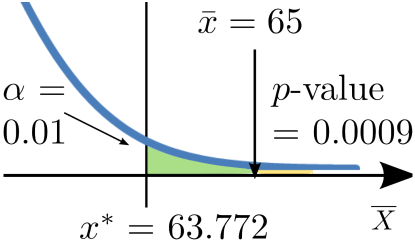
Example 1: graphical illustration

- Because $\bar{x} = 65$ falls in the rejection region $(63.722, \infty)$, we reject the null hypothesis.



Example 1: graphical illustration

- Because p-value $= 0.000899 < 0.01 = \alpha$, we reject the null hypothesis.



Road map

- Preparations.
- Testing population mean: variance known.
- **Testing population mean: variance unknown**.
- Testing population proportion.

Example 2

- An MBA program seldom admits applicants without a work experience longer than two years.
- To test whether the average work year of admitted students is above two years, 20 admitted applicants are randomly selected.
- Their work experiences prior to entering the program are recorded (in Sheet "Example 2" in "SDA-Fa14_11_testing2.xlsx.")
- The population is believed to be normal.

Example 2: hypothesis

- Suppose the one asking the question is a potential applicant with one year of work experience. He is **pessimistic** and will apply for the program **only if** the average work experience is proven to be **less** than two years.
- The hypothesis is

$$H_0 \colon \mu = 2$$
$$H_a \colon \mu < 2.$$

  - μ is the average work experience (in years) of all admitted applicants prior to entering the program.
  - To **encourage** him, we need to give him a strong evidence showing that his chance is high.

Example 2: hypothesis and test

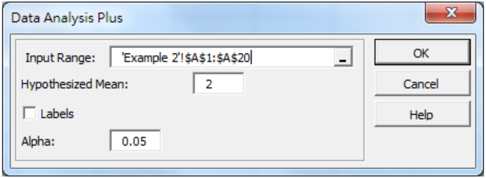
- Suppose he is **optimistic** and will not apply for the program **only if** the average work experience is proven to be **greater** than two.
- The hypothesis becomes

$$H_0 : \mu = 2$$
$$H_a : \mu > 2.$$

  - To **discourage** him, we need to give him a strong evidence showing that his chance is slim.
- Let's consider the optimistic candidate (and $H_a : \mu > 2$) first.
- Because the population variance is unknown and the population is normal, we may use the t test.

Example 2A: test

- In MS Excel, DAP → T-Test: Mean.



	A	B	C	D
1	t-Test: Mean			
2				
3				*Column 1*
4	Mean			2.5
5	Standard Deviation			1.3765
6	Hypothesized Mean			2
7	df			19
8	t Stat			1.6245
9	P(T<=t) one-tail			0.0604
10	t Critical one-tail			1.7291
11	P(T<=t) two-tail			0.1208
12	t Critical two-tail			2.093

- The one-tailed p-value is 0.0604.

Example 2A: test

- Alternatively, we may do the test step by step.
  - In Cell A21: $\bar{x} =$ `AVERAGE(A1:A20)` $= 2.5$.
  - In Cell A22: $s =$ `STDEV(A1:A20)` $= 1.376$.
  - In Cell A23: If H_0 is true and thus $\mu = 2$, the t statistic

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{2.5 - 2}{1.112/\sqrt{20}} = \texttt{(A21 - 2) / (A22 / SQRT(20))} = 1.6245.$$

  - In Cell A24: The p-value $=$ `TDIST(A23, 19, 1)` $= 0.0604$.
- In R, execute `t.test(x, alternative = "g", mu = 2)`, where `x` is the vector containing the sample data.

Example 2A: interpretation

- Conclusion:
  - For this one-tailed test, as p-value $= 0.0604 > 0.05 = \alpha$, we do not reject H_0.
  - There is **no strong evidence** showing that the average work experience is longer than two years.
  - The result is not strong enough to discourage the potential applicant, who has only one year of work experience.
- Decision:
  - The (optimistic) applicant **should** apply.

Example 2B – a pessimistic applicant

- Suppose the applicant is pessimistic and the hypothesis is

$$H_0: \mu = 2$$
$$H_a: \mu < 2.$$

  - The p-value will be $1 - 0.0604 = 0.9396$.[5]
  - We do not reject H_0 and cannot conclude that $\mu < 2$. There is no strong evidence to encourage him.
  - He **should not** apply.
- Note that when we write different alternative hypotheses, the final decision is different!
  - This happens if and only if in both cases we do not reject H_0.

[5] In R, execute `t.test(x, alternative = "l", mu = 2)`, where `x` is the vector containing the sample data.

Road map

- Preparations.
- Testing population mean: variance known.
- Testing population mean: variance unknown.
- **Testing population proportion**.

Testing the population proportion

- In many situations, we need to test the **population proportion**.
  - The defective rate or yield rate of a production system.
  - The proportion of people supporting a candidate.
  - The proportion of people supporting a policy.
  - The proportion of people viewing a product web page that will really buy the product (conversion rate).
- How to test the population proportion?
- Suppose we want to test the proportion of male users:
  - Let's label a male user by 1 and non-male users by 0.
  - Then the population proportion $p = \frac{\sum_{i=1}^{N} x_i}{N}$, the **population mean**.
  - A sample proportion $\hat{p} = \frac{\sum_{i=1}^{n} x_i}{n}$, the sample mean.
  - We may apply **the z test** to test population proportion.[6]
- Technical restrictions: $n \geq 30$, $n\hat{p} \geq 5$, and $n(1 - \hat{p}) \geq 5$.

[6] We may derive σ^2 from p for 0-1 data.

The hypotheses

- The population proportion is denoted as p.
- A two-tailed test for the population proportion is

$$H_0: p = p_0$$
$$H_a: p \neq p_0,$$

  where p_0 is the **hypothesized proportion**.
- In a one-tailed test, the alternative hypothesis may be either

$$H_a: p > p_0$$

  or

$$H_a: p < p_0.$$

Example 3

- In a factory, it seems to us that the defective rate of our product is too high. Ideally it should be below 1% but some workers believe that it is above 1%.
- If the defective rate is above 1%, we should fix the machine. Otherwise, we do not do anything.
- Let p be the defective rate, the hypothesis is

$$H_0 : p = 0.01$$
$$H_a : p > 0.01.$$

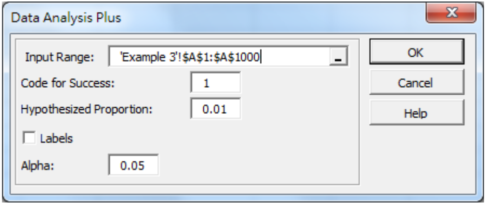
- When to adopt $H_a : p < 0.01$?

Example 3

- In several random production runs, we found that out of 1000 produced items, 14 of them are defective.
  - Sheet "Example 3" in "SDA-Fa14_11_testing2.xlsx."
  - The observed sample proportion $\hat{p} = 0.014$.
  - All the technical requirements are satisfied; $n = 1000$, $n\hat{p} = 14$, and $n(1 - \hat{p}) = 986$.
- Suppose the significance level is set of $\alpha = 0.05$, what is our conclusion?

Example 3: calculation

- In MS Excel, DAP → Z-Test: Proportion.[7]



	A	B	C	D
1	**z-Test: Proportion**			
2				
3				*Column 1*
4	Sample Proportion			0.014
5	Observations			1000
6	Hypothesized Proportion			0.01
7	z Stat			1.2713
8	P(Z<=z) one-tail			0.1018
9	z Critical one-tail			1.6449
10	P(Z<=z) two-tail			0.2036
11	z Critical two-tail			1.96

- The one-tailed p-value is 0.1018.

[7] In R, execute `prop.test(x = 14, n = 1000, p = 0.01, alternative = "g", correct = FALSE)`.

Example 3: conclusion and decision

- Conclusion:
  - For this one-tailed test, as p-value $= 0.1018 > 0.05 = \alpha$, we do not reject H_0.
  - There is **no strong evidence** showing that the defective rate is higher than 1%.
- Decision:
  - We should not try to fix the machine.